9th NOAA TBPG Workshop
Kansas City, MO
April 10-11, 2018

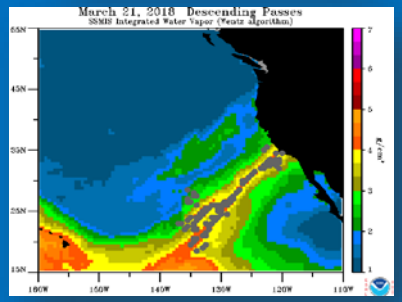
Roundup Presentation

Hydrometeorology Testbed (HMT)

James Nelson (NWS/WPC), Lisa Darby (OAR/PSD),
Allen White (OAR/PSD)

Hydrometeorology Testbed
Co-managed by NWS/WPC and OAR/PSD

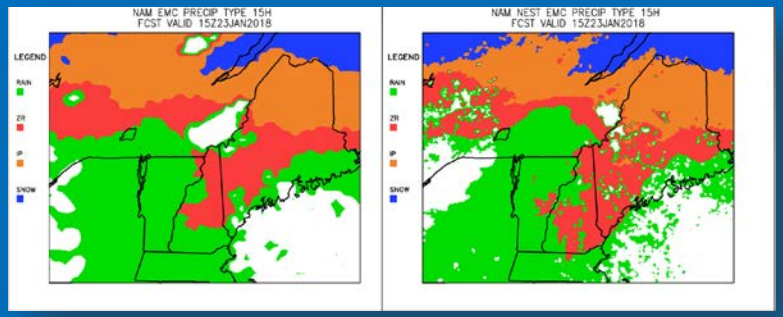
- **Vision**
  - Communities that are more resilient to the impacts of precipitation extremes on lives, property, water supply and ecosystems.
- **Overarching Question**
  - Can forecasts of precipitation extremes be improved through an increased predictive understanding of underlying physical processes derived from enhanced regional observations and models?
- **Mission**
  - **To accelerate the development, prototyping, and use** of advanced hydrometeorological observations and models to improve our physical understanding and representation of precipitation and surface processes, leading to improved monitoring and prediction of extreme events
  - **To foster infusion of these advances** into NOAA services, and to provide customized information for local and regional decision making
  - **To support the broader needs** for 21st Century precipitation information for flood and water resources management and improved ecosystem services



HMT at WPC

8th Annual Winter Weather Experiment (WWE) at WPC

- **Focus on forecasting precipitation type**
  - Explore multiple microphysical and probabilistic precipitation type methodologies
  - Explore the applications of fuzzy clustering to forecasting winter weather
  - Enhance collaboration among NCEP centers, WFOs, and NOAA research labs
- **Experiment operations**
  - Weekly experiments from Nov. 14th 2017 – Mar. 9th 2018
    - Forecast Tuesdays and Verification Wednesdays
    - 2 weeks of in-house experiments at WPC
- **Products from the following models were included in the experiment**
  - NAM/Nest, PWPF ensemble, HRRRX, HRRRE, HREFv2, FV3-GFS, NBMv3



NCEP Dominant precipitation type

From a presentation given by Eric Aligo, Brad Ferrier, Mike Bodner

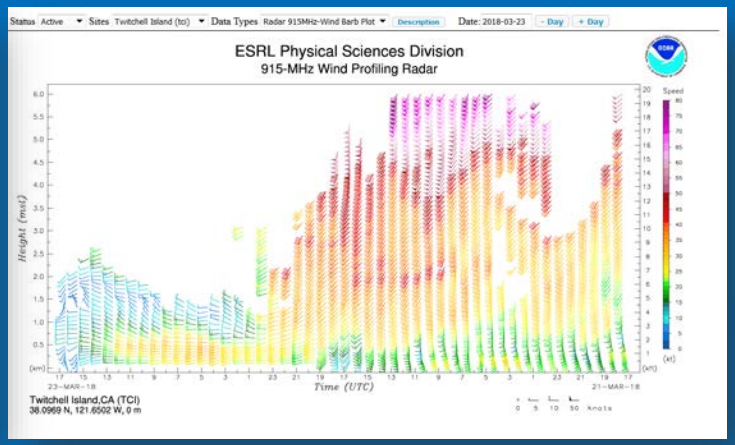
HMT at WPC

Flash Flood and Intense Rainfall Experiment (FFaIR) at WPC

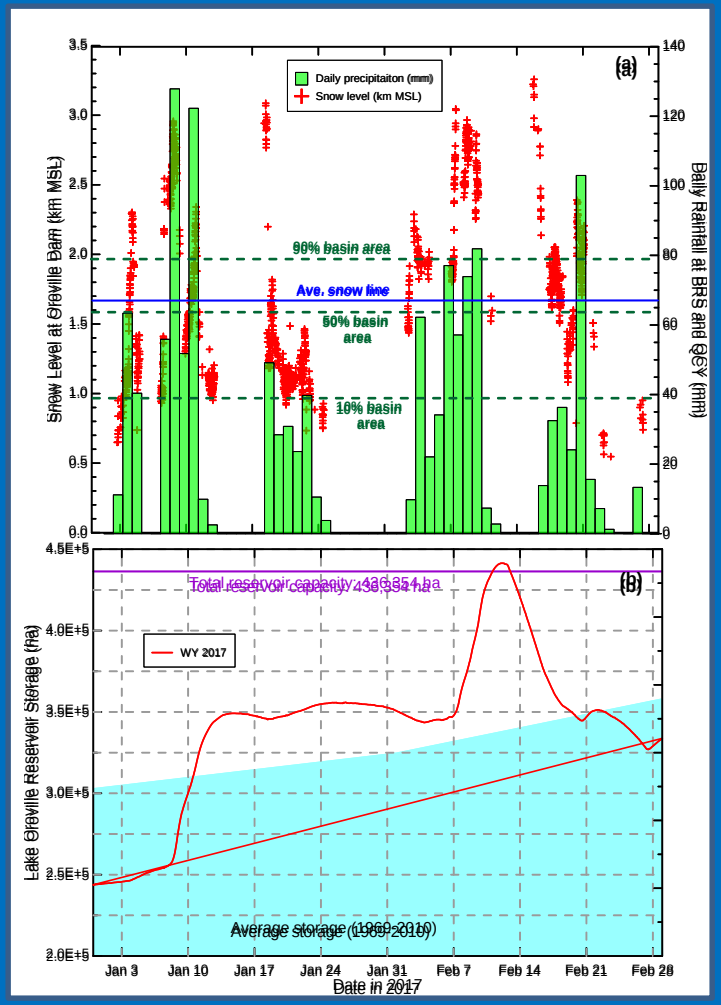
- **Focus on high-resolution guidance to improve flash flood forecasts**
  - Maximize the utility of high-resolution convection-allowing models and ensembles for short-term forecasts
  - Maximize the utility of high-resolution convection-allowing deterministic models for Days 2 & 3
  - Identify the most effective forms and proper usage of available hydrologic guidance for assessing flood risk
  - Explore proposed changes to WPC's operational Excessive Rainfall Outlook by evaluating an experimental "first guess" recommender for Day 2 & 3
  - Enhance collaboration among NCEP centers, WFOs, and NOAA research labs
- **Experiment operations**
  - 4 Weekly in-house experiments in June and July
- **Products from the following models were included in the experiment**
  - NAM, NAM/Nest, Flash Flood guidance, HRRR, NMMB ARW WRF-NSSL, NWM, Precipitation Recurrence Data (Atlas 14), HRRRv3, HRRR-TLE, HREFv2, WRF-ARW SSEFX, FV3-CAPS, UM, FV3-GFS, FV3-GFDL, NBMv3

HMT at PSD

- At the request of CA-DWR and in response to the Oroville Dam flood mitigation crisis that occurred in February 2017, PSD installed two additional atmospheric river observatories to monitor atmospheric river moisture transport into the Feather River Basin:
  - Twichell Island, at the east end of the San Francisco Bay Delta to monitor atmospheric river moisture transport through this prominent gap in coastal terrain
  - Thermalito Power Station just upwind of the Oroville Dam



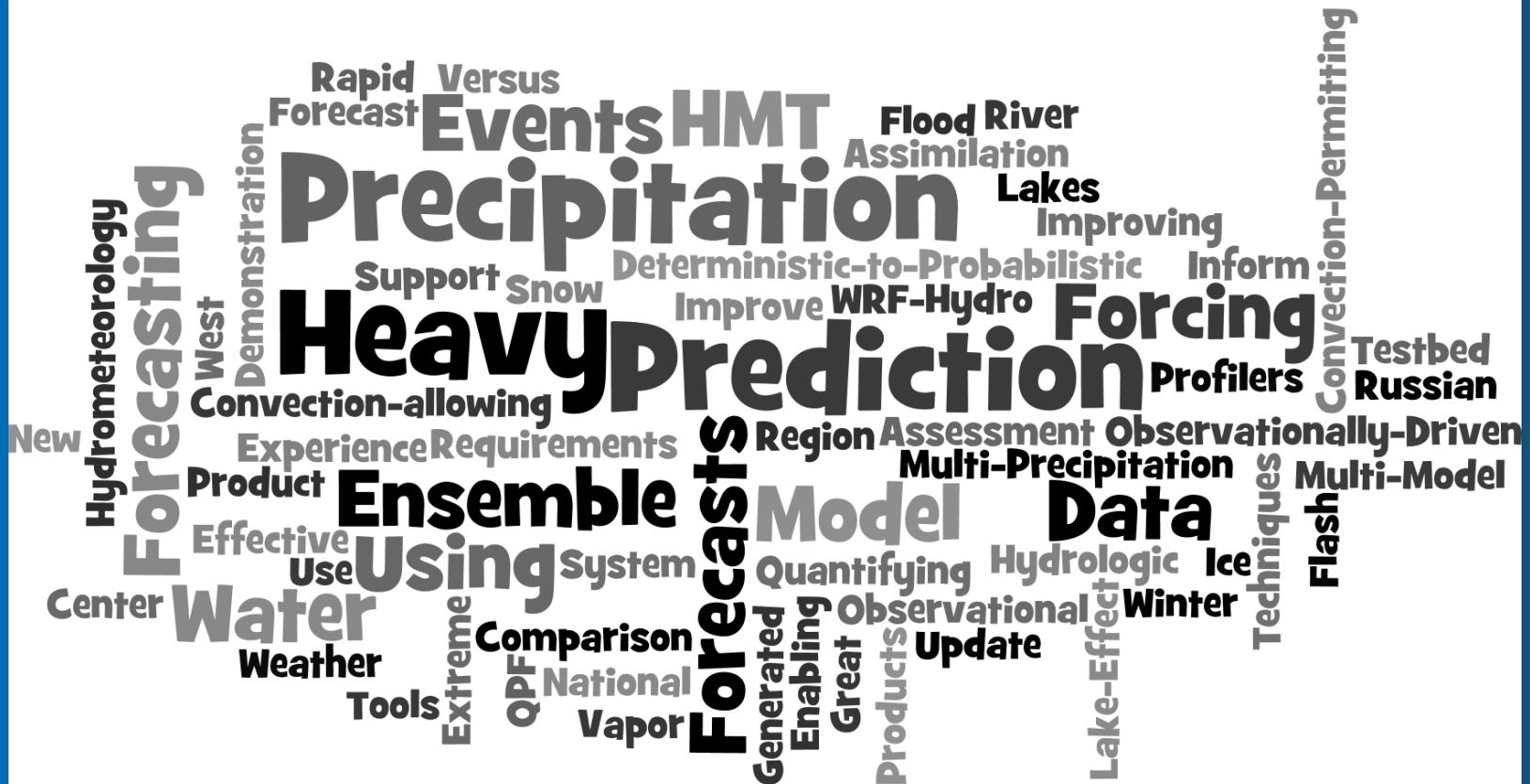
HMT at PSD



Scientists at PSD are analyzing precipitation events preceding the Oroville Dam crisis – assessing rain vs. snow levels over the feather River Basin and the impact on reservoir levels

- Top: Daily accumulated precipitation (light green bars) averaged from 3 gauges in the Feather River Basin for January and February 2017 along with
- Snow levels (km MSL) observed by the snow-level radar located at the base of Oroville Dam (red plus symbols).
- Horizontal lines denote the average snow line in the watershed and three basin area elevation levels (% of basin area; dashed dark green lines).
- Bottom: Hourly storage level (ha) in Lake Oroville during the same period as in top plot (red curve); monthly average storage level (cyan shaded area); and reservoir capacity (purple line). Precipitation gauge and reservoir storage data were acquired through CDEC.

HMT OWAQ-funded projects

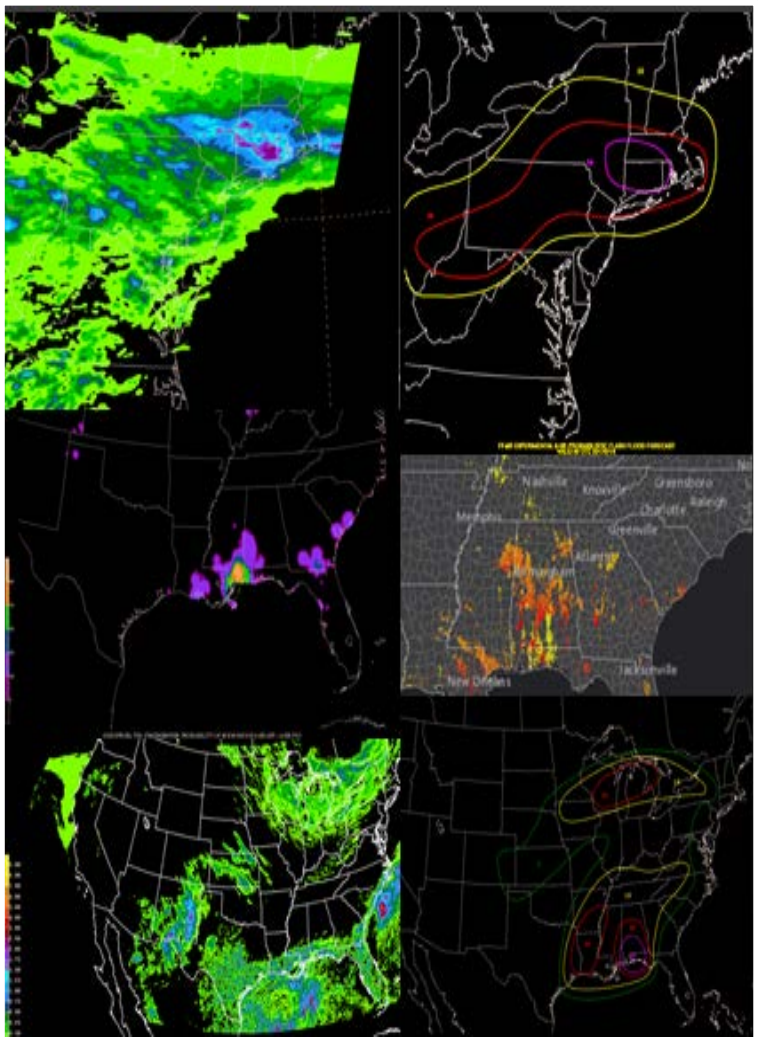


FY18 Highlights

WPC-HMT – 2017 FFaIR Experiment

GOALS AND OBJECTIVES

- Identify ways to maximize the utility of high resolution convection-allowing models and ensembles for short-term flash flood forecasts
- Evaluate the utility of high resolution convection-allowing deterministic models and hydrologic models for flood forecasts at longer time ranges (Day 2-3)
- Identify effective forms and proper usage of available hydrologic guidance for the assessment of flood risk
- Explore proposed changes to WPC's operational Excessive Rainfall Outlook (ERO)
- Enhance collaboration between the operational forecasting, research, and academic communities on the forecast challenges associated with short-term flash flood forecasting



FY18 Highlights

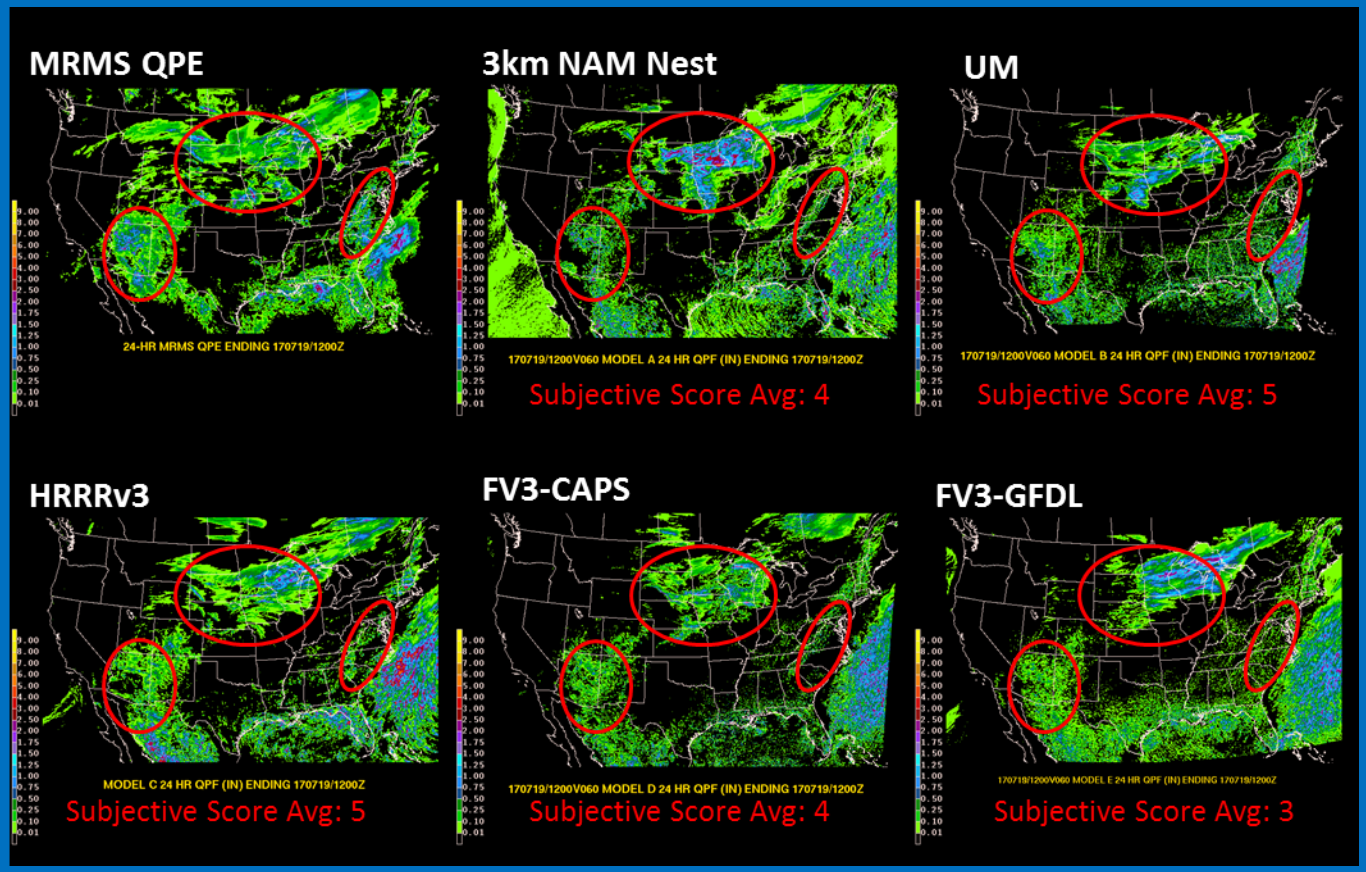
WPC-HMT – 2017 FFaIR Experiment

Example Day

Subjective Verification for Day 2 (36-60 hr forecasts*) QPF Models.

Valid 12Z July 18 -- 12Z July 19, 2017.

Monsoonal moisture in the Southwest, Northern Plains moisture, scattered Southeast moisture.

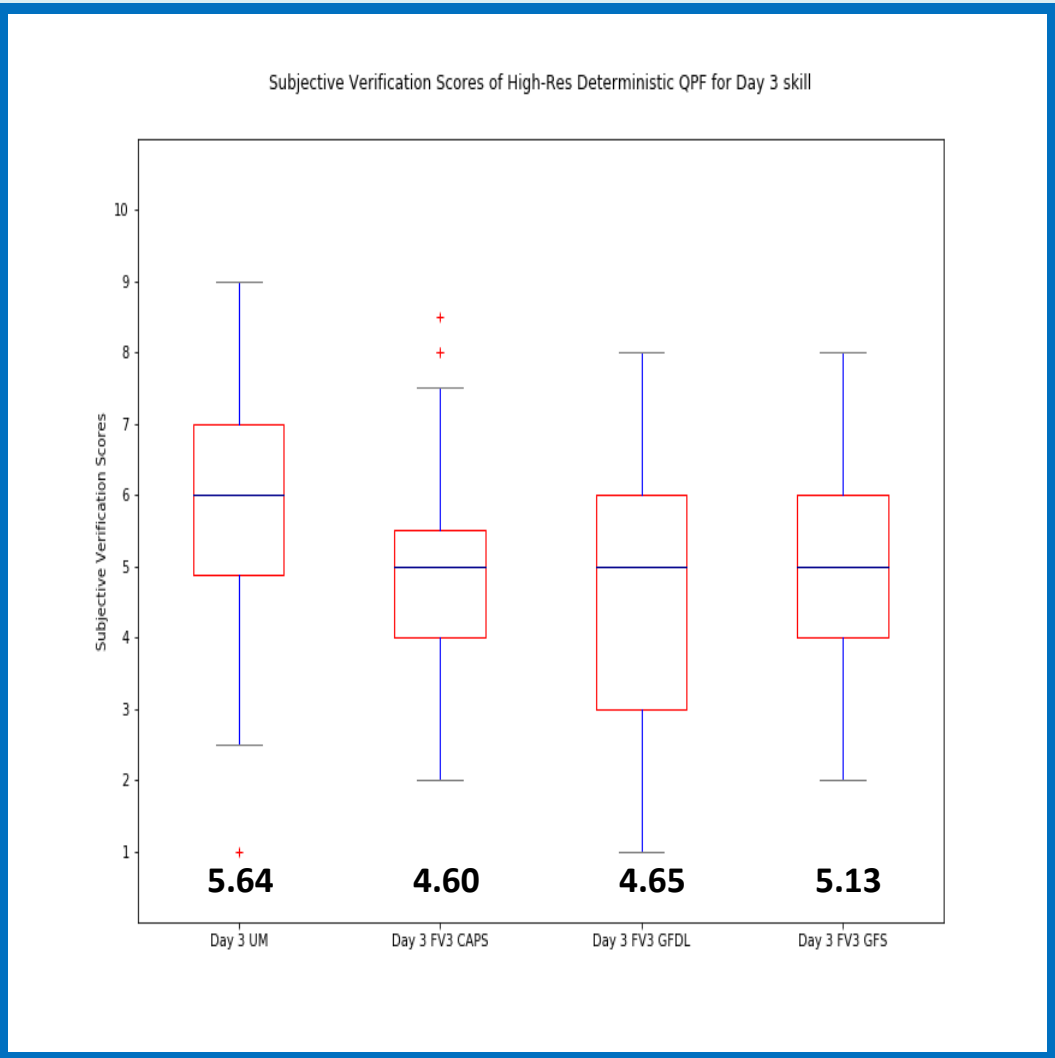


Day 3 Model QPF Subjective Evaluation

NOTE: Due to data availability, the number of scoring opportunities differed among models.

General comments:

- Subjectively, UM had highest average score.
- Lower resolution FV3-GFS had higher average subjective score than high-res FV3-CAPS and FV3-GFDL.
- Participants commented the FV3-GFS did well with spatial representation of precipitation but often struggled with higher amounts over a half inch.
- FV3-CAPS Day 3 subjective score average was higher than Day 2. Saw several examples where Day 3 forecasts scored better than Day 2.

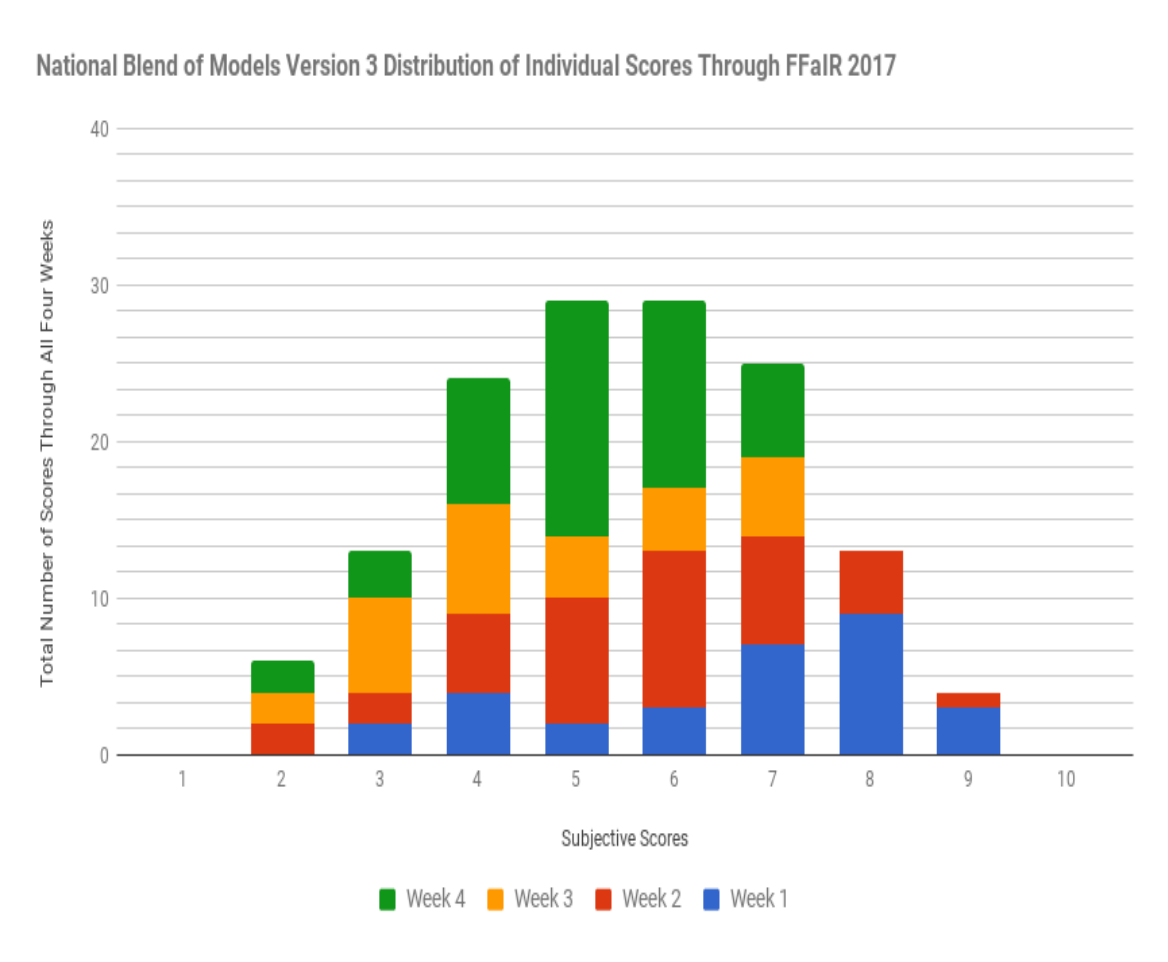


National Blend of Models Version 3 6 HR QPF

NBMv3
Average Subjective Score: 5.41

General comments:

- Highest scores were seen in Week 1 and Week 2 (blue and red on the chart).
- Week 1 especially had more synoptic-scale systems. Weeks 3 and 4 were dominated by more mesoscale MCS activity and Southwest monsoon moisture.
- Most common criticism from participants was a low bias in the QPF amounts.
- Did better with spatial coverage, however at times was too broad with lighter amounts.



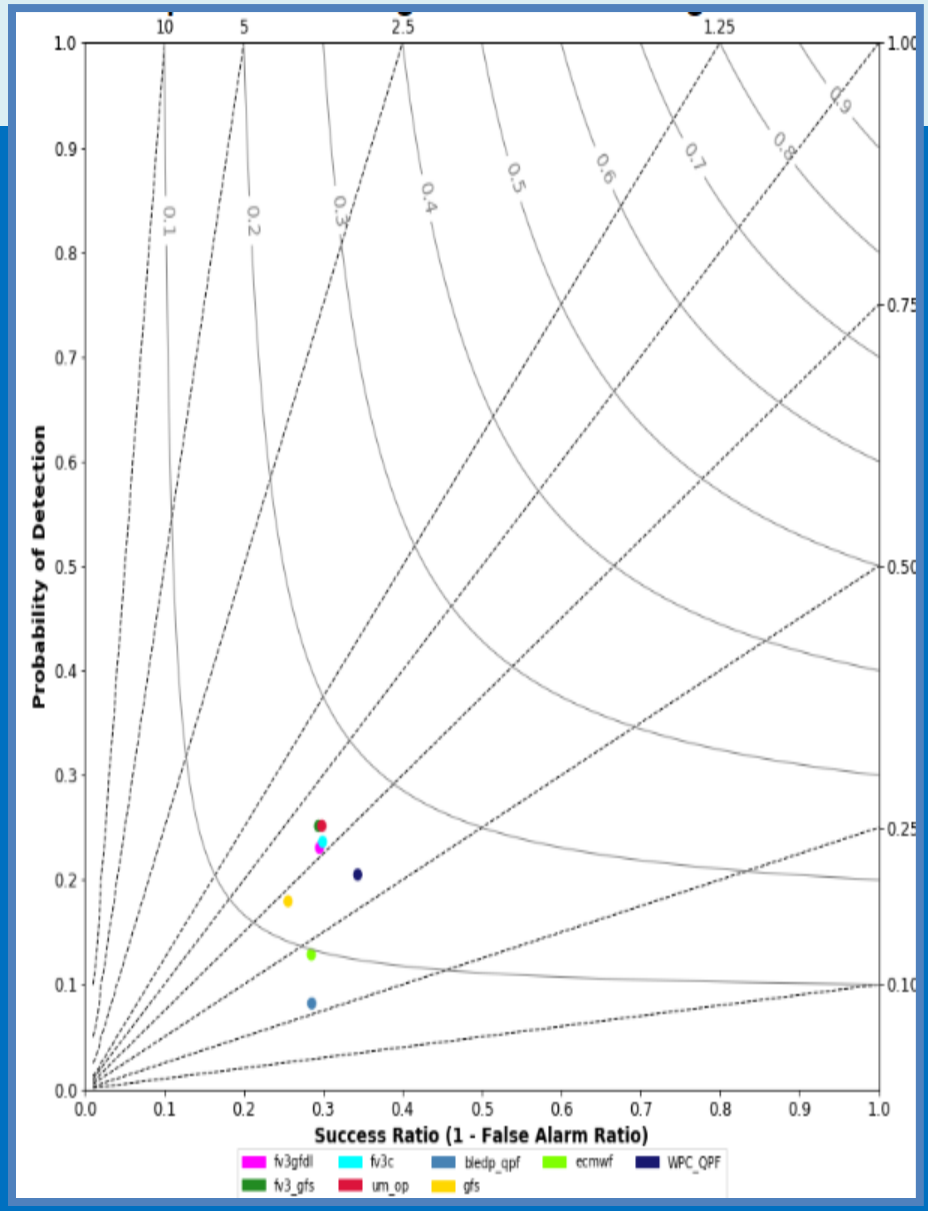
OBJECTIVE VERIFICATION FOR EXPERIMENTAL GUIDANCE

DAY 3, 0.5” QPF THRESHOLD

General comments:

- FV3-GFS 0.5” QPF scored highest on overall MODE analysis (not shown) most frequently correctly forecasting large continuous areas of heavy QPF, with some metrics outperforming 60 hour models
- The UM again had the highest MODE CSI value
- The WPC forecast did not rank favorably at this threshold

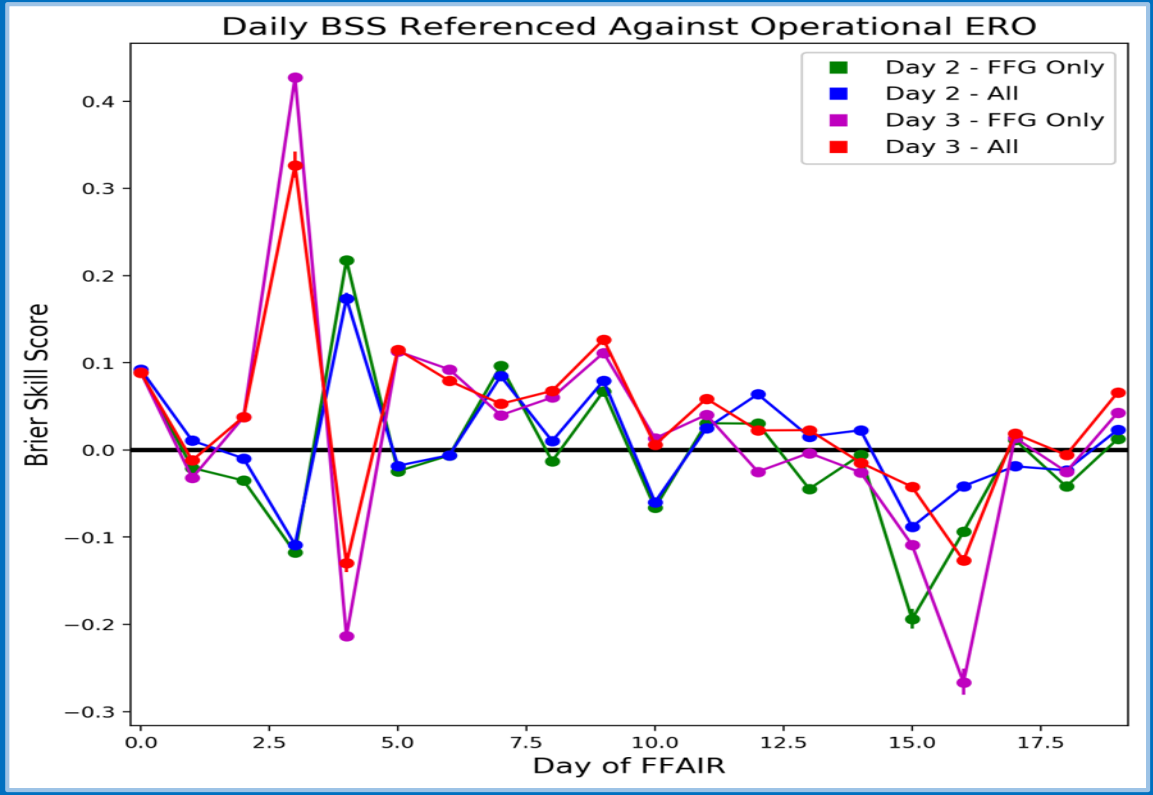
CSI Mean	FBIAS Mean	Model
0.161763	0.868279	UM
0.158819	0.813155	FV3-GFS
0.156749	0.773107	FV3-CAPS
0.149191	0.759035	FV3-GFDL
0.143407	0.610742	WPC



Daily Time-series of Brier Skill Score

- To assess probabilistic skill, BSS for the FFaIR ERO is referenced against the operational ERO for each day during FFaIR.

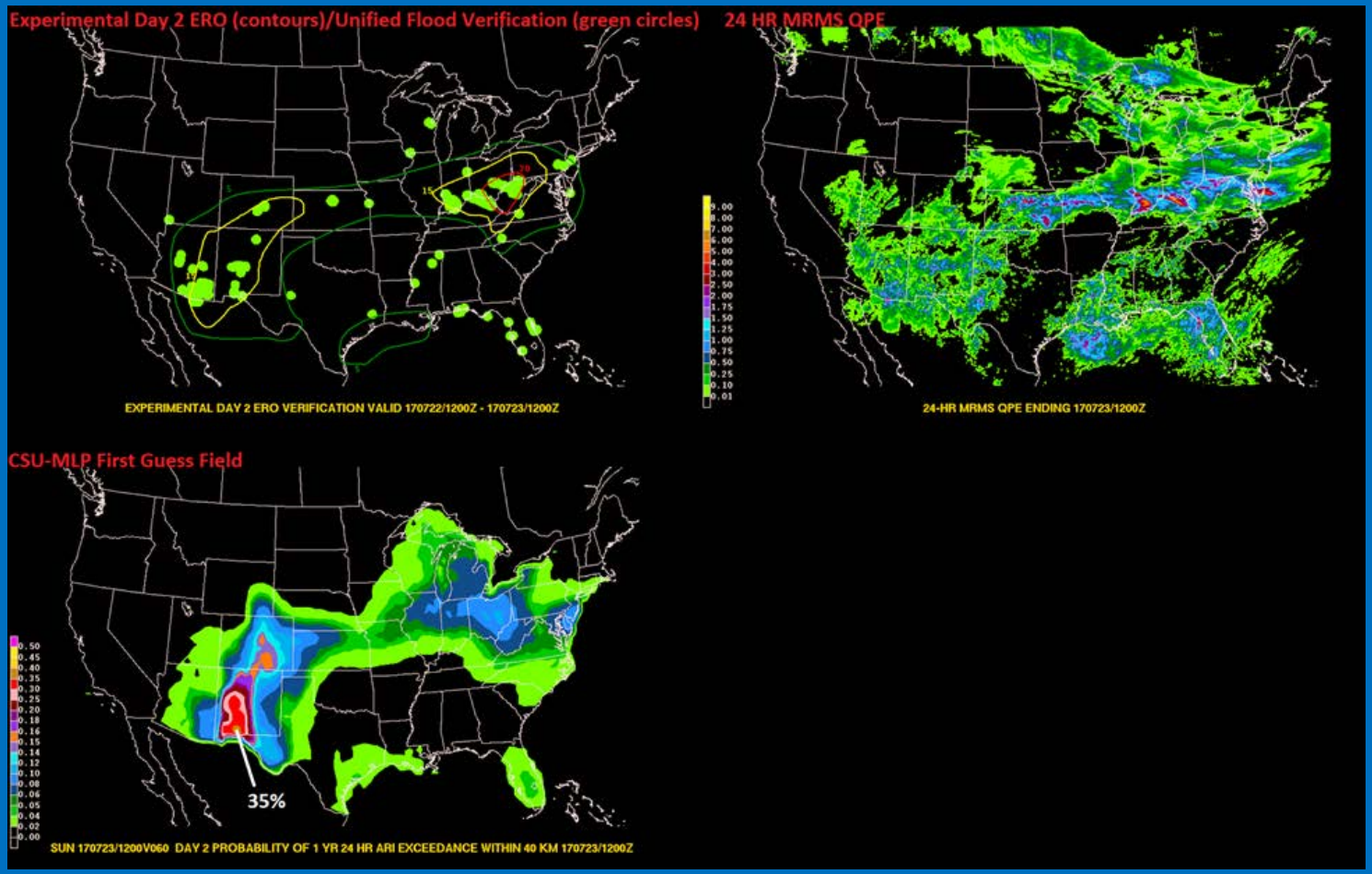
- The FFaIR ERO exhibits lower probabilistic error (BSS > 0) more frequently on weeks 1 and 2, when greater large-scale forcing is present.



CSU-MLP First Guess ERO

The **CSU Machine-Learning First Guess field** for the ERO showed great potential and was scored well by participants.

It was recommended that the CSU developers work to reduce recurring biases and continue to refine the tool and reintroduce it into the testbed 2018 FFaIR for further evaluation.



FY18 Transition Metrics

HMT: Flash Flood and Intense Rainfall Experiment (FFaIR)

Summer 2017

Major Tests Conducted	Transitioned to Operations (RL9)	Recommended for Transition to Operations (RL9)	Advanced To Experimental Testing (RL8)	Further Demonstration/ Development (RL 5-7)	Rejected For Further Testing
CSU-MLP ARI/GEFS-R ERO First Guess Field				X	
EAS Variable Ensemble Probabilistic Filter Technique		X			
Ensemble Blended Mean Post-Processing		X			
HRRR-TLE Probability of QPF on Saturated Soil		X			
National Water Model: Time to High Flow				X	
National Water Model: Ponded Water					X
National Water Model: High Flow Probability				X	

FY18 Highlights

WPC-HMT – 2017 FFaIR Experiment

2017 FFaIR Recommendation

- It continues to be strongly recommended that water models and precipitation models continue to experiment with fusion (not only in model forcings, but also in visualization) to create products best suited for flood forecasters with issuance deadlines. It continues to be the desire of forecasters to be able to overlay the model QPF onto the hydrologic response data, both deterministically and probabilistically, with animating time steps for full situational awareness and improved prediction and messaging.

FY18 Highlights

WPC-HMT – 2017/18 Winter Weather Experiment

FY18 Highlights

WPC-HMT – 2017/18 Winter Weather Experiment

- 2017/18 WWE Goals
  - Explore the testing of multiple microphysical and probabilistic precipitation type methodologies to determine which methods best enhance the forecast process.
  - Explore the applications of Fuzzy Clustering to the prediction of winter weather
  - Enhance collaboration among NCEP centers, WFOs, and NOAA research labs on winter weather forecast challenges.

FY18 Highlights

WPC-HMT – 2017/18 Winter Weather Experiment

2017/18 WWE Preliminary Results - PTYPE

- NAM - The identification of areas of snow was good in most cases using this methodology, however rain and freezing rain were too often not forecast.

- NBM - The majority of the NBM probabilities of precipitation type were useful, and forecasters were most impressed out at Day 3. At shorter time scales, the probabilities tended to be too low and missed snow, freezing rain and sleet in areas over which it was observed. This was particularly the case in areas of complex terrain and near coastlines.

FY18 Highlights

WPC-HMT – 2017/18 Winter Weather Experiment

2017/18 WWE Preliminary Results - PTYPE

- FV3 – There still was a tendency to not forecast enough sleet and freezing rain in precipitation type transition zones.

- Fuzzy Clustering - There was consensus among experiment participants that the use of clustering techniques shows promise in improving the forecast process by providing information on ensemble sensitivity.

❑ Recommendation - the development of training material to enable increased use among forecasters as well as exploring ways to more intuitively visualize and interrogate the data in the display.

FY18 Transition Metrics

HMT: Winter Weather Experiment

Winter 2017-2018

Major Tests Conducted	Transitioned to Operations (RL9)	Recommended for Transition to Operations (RL9)	Advanced To Experimental Testing (RL8)	Further Demonstration/ Development (RL 5-7)	Rejected For Further Testing
NBM Conditional Prob Snow				X (It is oper in NBM)	
NBM Conditional Prob Ice				X (It is oper in NBM)	
NBM Prob of Precipitation Type				X (It is oper in NBM)	
ESRL/GSD Variable Density SLR for application in the HRRR/HRRRE			X		
HREF EAS method for prob of PTYPE			X		
Stony Brook Fuzzy Clustering for PTYPE				X	

FY18 Needs

WPC-HMT

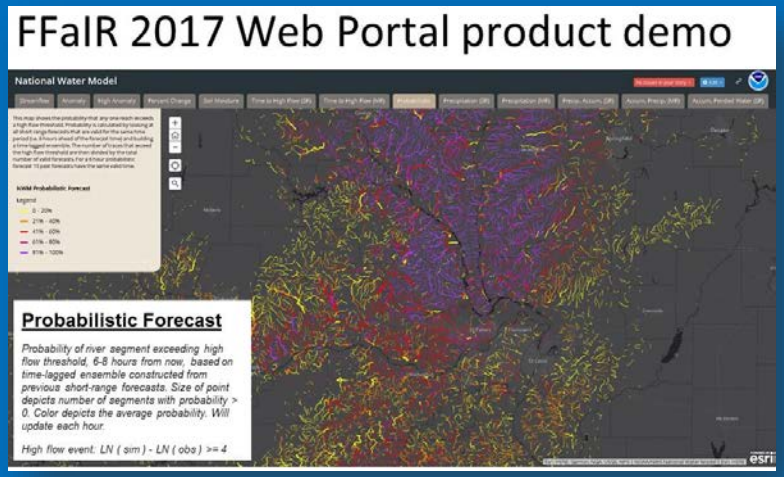
WPC-HMT Needs

- Reliable/Valid PTYPE Verification Dataset
- Questions rose to the validity of all verification data sets including ASOS stations, NOHRSC snowfall analyses, the reporting time of local storm reports, and the thermal profile in the RAP analyses.
- AWIPS Transition
- WPC-HMT needs to accelerate the transition to AWIPS2. This will make transition of datasets and tools to WFOs/RFCs more efficient.

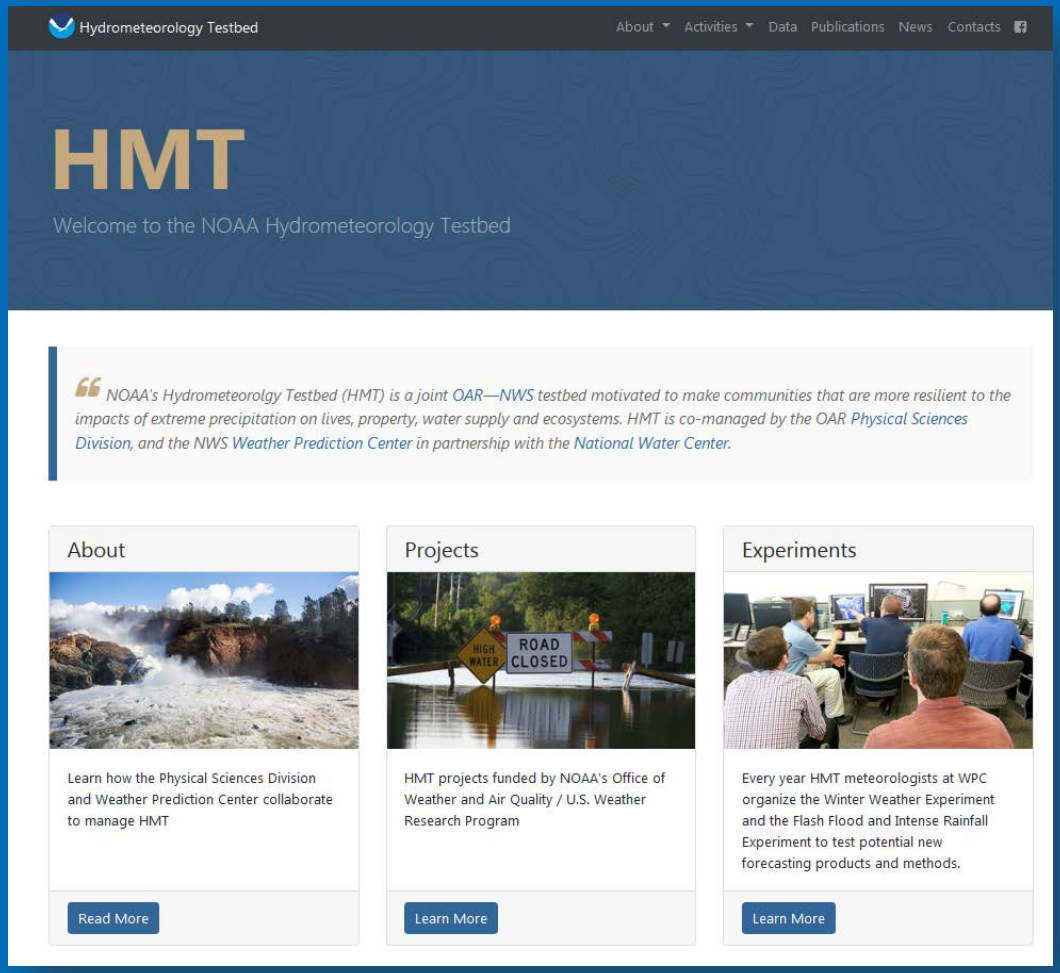
FY17 Highlights

HMT

- **FY17 Highlight – Demonstration of Advanced Ensemble Prediction Services for NWS Hydrometeorological Forecast Operations (Mahoney & Gochis)**
  - *One of the "OAR Top Five" USWRP projects for transition to operations*
  - Demonstrate hydrologic time-lagged ensemble techniques for the National Water Model (NWM)
  - Provide guidance on design of the NWM medium-range ensemble configuration
  - Create and demonstrate probabilistic verification and visualization techniques
  - FFaIR forecaster feedback was critical to improving the products developed by this OWAQ-funded project
  - Time to High Flow and High Flow Probability products are being tested at the NWC



HMT has a new web site!



- Combined content from ESRL Physical Sciences Division and NCEP Weather Prediction Center
- HMT has evolved over the years – needed a new website to reflect the current HMT model
- More content to come

Questions

HMT

- **Additional Information**
  - **Jim Nelson (James.A.Nelson@noaa.gov)**
  - **Lisa Darby (Lisa.Darby@noaa.gov)**
  - **https://hmt.noaa.gov/**

Backup Slides

FY18 Highlights

WPC-HMT – 2017 FFaIR Experiment

- 2017 FFaIR Results
  - The CSU Machine-Learning First Guess field for the ERO showed great potential and was scored well by participants.
  - The ensemble QPF blended mean was favorably scored by participants and was generally successful in increasing QPF magnitude of the composite mean and decreasing the magnitude of the probability matched mean.
  - The Time-Lagged HRRR (HRRR-TLE) QPF on saturated soil product was a useful tool for visualizing high probabilities of precipitation falling on soil that is already near or at saturation.
  - The Experimental Time to High Flow product from the National Water Model (NWM) has high potential for use by forecasters in assessing flood risk.
  - The Experimental Ponded Water from the NWM is not recommended for operations at this time.
  - The Experimental High Flow Probability product from the NWM is a welcome first step into probabilistic hydrologic guidance.

FY18 Highlights

WPC-HMT – 2017 FFaIR Experiment

- 2017 FFaIR Results
  - Testing Day 2 and Day 3 high resolution deterministic QPF was a primary goal of the 2017 FFaIR Experiment, and results were mixed. Common weaknesses in convective allowing models, including difficulty capturing events in more weakly-forced synoptic regimes and a low bias in the QPF magnitude and spatial extent, propagate from Day 1 into Days 2 and 3. There is slight degradation in skill from Day 2 to Day 3. However, participants found value in the higher resolution models, especially reflectivity fields, and can apply known biases to the forecast process. It is recommended that developers continue work on the schemes that improve QPF and reflectivity fields in high-resolution models so they may further benefit flood forecasting at longer time ranges.
  - The NBMv3 QPF had a magnitudinal low bias for the majority of the cases. As with the deterministic models, synoptic systems were forecast more optimally than mesoscale features, such as Southwest monsoon moisture or MCS activity. The overall shape and spatial coverage was often quite good and gave an indication that there would be precipitation in areas that received rain.